

Modelos predictivos de caries dental basados en Big Data y aprendizaje automático: revisión narrativa

Predictive models of dental caries based on Big Data and machine learning: a narrative review

Noelia Amaray Velastegui Almeida^I

noelia.velastegui@ug.edu.ec

<https://orcid.org/0009-0004-1745-3986>

Gonzalo Paúl Rodríguez Galarza^{II}

gorodriguezga@uide.edu.ec

<https://orcid.org/0009-0009-1449-4041>

Correspondencia: noelia.velastegui@ug.edu.ec

Artículo de Revisión

Recibido: 28 de enero del 2026

Aceptado: 11 de marzo del 2026

Publicado: 25 de marzo del 2026

- I. Estudiante de la Carrera de Odontología, Universidad de Guayaquil, Guayaquil, Ecuador.
- II. Ingeniero en Electrónica y Telecomunicaciones. Posgradista, Universidad Internacional del Ecuador, Quito, Ecuador.

Cómo citar este artículo:

Velastegui, N., & Rodríguez, G. (2026). Modelos predictivos de caries dental basados en Big Data y aprendizaje automático: revisión narrativa. *Revista Colincing de Estudios Multidisciplinarios*, 2(1), e15. <https://doi.org/10.61347/rcem.v2i1.e15>

Copyright:

Derechos de autor 2026 Noelia Amaray Velastegui Almeida, Gonzalo Paúl Rodríguez Galarza.



Esta obra está bajo una licencia internacional Creative Commons Atribución-NoComercial 4.0 (CC BY-NC 4.0)

Resumen: La caries dental es la enfermedad bucal más prevalente a nivel mundial, afectando aproximadamente a 3,5 mil millones de personas y generando un impacto significativo en la calidad de vida y en los sistemas de salud. Ante esta problemática, los modelos predictivos han emergido como herramientas clave para la detección temprana y la prevención; sin embargo, muchos presentan limitaciones metodológicas, como alto riesgo de sesgo y baja aplicabilidad clínica. En este contexto, el avance del Big Data, la inteligencia artificial y el aprendizaje automático ha impulsado el desarrollo de modelos más precisos y eficientes. La presente revisión narrativa tuvo como objetivo analizar la evidencia científica sobre la aplicación de estas tecnologías en la predicción del riesgo de caries dental, incluyendo estudios publicados entre 2015 y 2025 en bases de datos como Scopus, PubMed y Web of Science. Los resultados evidenciaron que las redes neuronales convolucionales, las máquinas de vectores de soporte y las redes neuronales artificiales son los algoritmos más utilizados, destacando por su alto rendimiento diagnóstico, con valores elevados de precisión y área bajo la curva. Las imágenes radiográficas fueron el tipo de dato más empleado, seguidas de datos clínicos y modelos mixtos. No obstante, se identificaron limitaciones importantes, como la heterogeneidad de los datos, la falta de estandarización de métricas y la escasa validación externa. En conclusión, los modelos basados en inteligencia artificial muestran un alto potencial para mejorar la predicción de caries dental, aunque es necesario fortalecer el rigor metodológico y su validación clínica para favorecer su implementación en la práctica odontológica.

Palabras clave: Aprendizaje automático, Big Data, caries dental, inteligencia artificial, modelos predictivos.

Abstract: *Dental caries is the most prevalent oral disease worldwide, affecting approximately 3.5 billion people and generating a significant impact on quality of life and healthcare systems. In response to this challenge, predictive models have emerged as key tools for early detection and prevention; however, many of them present methodological limitations, such as a high risk of bias and limited clinical applicability. In this context, advances in Big Data, artificial intelligence, and machine learning have driven the development of more accurate and efficient models. This narrative review aimed to analyze the scientific evidence on the application of these technologies in predicting the risk of dental caries, including studies published between 2015 and 2025 in databases such as Scopus, PubMed, and Web of Science. The results showed that convolutional neural networks, support vector machines, and artificial neural networks are the most used algorithms, standing out for their high diagnostic performance, with elevated values of accuracy and area under the curve. Radiographic images were the most frequently used type of data, followed by clinical data and mixed models. Nevertheless, important limitations were identified, including data heterogeneity, lack of standardized metrics, and limited external validation. In conclusion, artificial intelligence-based models show high potential to improve the prediction of dental caries, although strengthening methodological rigor and clinical validation is necessary to support their implementation in dental practice.*

Keywords: *Machine learning, Big Data, dental caries, artificial intelligence, predictive models.*

Introducción

La caries dental es una de las enfermedades infecciosas más prevalentes a nivel mundial, afectando aproximadamente a 3,5 mil millones de personas y constituyéndose en la afección bucal más frecuente (Shah, 2023). Su prevalencia ha mostrado una tendencia creciente, especialmente en países en desarrollo, como consecuencia de la adopción de dietas cariogénicas, la limitada exposición a fluoruros y las barreras económicas que dificultan el acceso a servicios odontológicos.

Asimismo, la caries dental es considerada la enfermedad no transmisible más común a nivel global, causada principalmente por el consumo excesivo de azúcares libres, lo que puede derivar en dolor, dificultad para la alimentación y pérdida de piezas dentales; en consecuencia, afecta significativamente la calidad de vida de los individuos y genera elevados costos para los sistemas de salud (World Health Organization [WHO], 2025).

El estudio de Wang et al. (2025) destaca que la elevada prevalencia y el impacto de la caries dental en niños y adolescentes evidencian la necesidad de implementar modelos predictivos que permitan su detección temprana y prevención. Estos modelos resultan fundamentales para identificar a los individuos con mayor riesgo de desarrollar la enfermedad; sin embargo, muchos de los modelos existentes presentan un alto riesgo de sesgo, asociado principalmente al uso de métodos estadísticos inadecuados y a la falta de transparencia en la presentación de los resultados.

Estas limitaciones reducen su aplicabilidad clínica, lo que pone de manifiesto la necesidad de fortalecer el rigor metodológico y promover validaciones externas en futuras investigaciones. De manera similar, Havsed et al. (2023) señalan que los modelos predictivos de caries desempeñan un papel clave en el diseño de intervenciones preventivas individualizadas y en la determinación de la periodicidad de la atención odontológica, ya que permiten estimar la probabilidad de aparición de nuevas lesiones a lo largo del tiempo.

No obstante, la mayoría de estos modelos presenta un rendimiento predictivo modesto y un elevado riesgo de sesgo, debido a deficiencias metodológicas como el número insuficiente de eventos analizados y la selección inadecuada de variables predictoras. Estas limitaciones generan estimaciones inciertas que reducen su utilidad clínica, lo que subraya la necesidad de mejorar los estándares metodológicos e incorporar nuevos predictores en el desarrollo de futuros modelos.

En este contexto, los avances tecnológicos han impulsado la incorporación de herramientas de inteligencia artificial (IA) y aprendizaje automático (machine learning, ML) en el campo de la odontología. El estudio de Ding et al. (2023) señala que la IA constituye un componente central de la denominada cuarta revolución industrial aplicada a la odontología, cuyo desarrollo se fundamenta en la disponibilidad de grandes volúmenes de datos (Big Data), el incremento de la capacidad computacional y el perfeccionamiento de algoritmos avanzados.

Estas tecnologías permiten que los sistemas informáticos aprendan a partir de datos clínicos, imágenes y otras fuentes de información, contribuyendo al diagnóstico, la planificación del tratamiento y la toma de decisiones clínicas. En particular, su aplicación en el análisis de imágenes dentales ha demostrado mejorar la eficiencia y la precisión diagnóstica.

De igual forma, Wang et al. (2025) destacan que el aprendizaje automático está transformando la práctica odontológica al mejorar la precisión y eficiencia en el diagnóstico, tratamiento y gestión de pacientes, mediante el uso de modelos avanzados como las redes neuronales artificiales (ANN), las redes neuronales convolucionales (CNN) y las máquinas de vectores de soporte (SVM).

Estas técnicas permiten analizar grandes volúmenes de datos clínicos y radiográficos, incluyendo radiografías y escaneos digitales, lo que facilita diagnósticos más precisos y el desarrollo de planes de tratamiento personalizados, superando diversas limitaciones de los métodos tradicionales de evaluación del riesgo de caries.

En este contexto, la presente revisión bibliográfica tiene como objetivo analizar y sintetizar de manera crítica la evidencia científica disponible sobre la aplicación de técnicas de Big Data y algoritmos de aprendizaje automático en la predicción del riesgo de caries dental en diversas poblaciones.

Específicamente, se busca identificar los tipos de datos utilizados, los algoritmos más empleados y las principales ventajas y limitaciones reportadas en la literatura, con el fin de proporcionar una visión integral del estado actual del conocimiento en este campo.

Este trabajo se fundamenta en las siguientes preguntas de investigación, formuladas para orientar la síntesis y el análisis crítico de la evidencia disponible: ¿Qué tipos de datos se utilizan para entrenar modelos de predicción de caries dental?; ¿Qué algoritmos de aprendizaje automático han demostrado mayor efectividad en la predicción de caries dental?; ¿Cuáles son las ventajas y limitaciones de estas herramientas en comparación con los métodos tradicionales de evaluación del riesgo de caries?; y ¿Qué tendencias y perspectivas existen para la implementación clínica de estas tecnologías en el ámbito odontológico?

Esta revisión bibliográfica busca contribuir al conocimiento científico actual mediante la integración crítica de la evidencia disponible sobre el uso de Big Data y aprendizaje automático en la predicción del riesgo de caries dental, destacando tanto los avances metodológicos como las limitaciones existentes en su aplicación clínica, con el propósito de orientar futuras investigaciones y promover el desarrollo de herramientas tecnológicas que fortalezcan la prevención, el diagnóstico temprano y la toma de decisiones en la práctica odontológica.

Metodología

La presente investigación se desarrolló mediante una revisión narrativa orientada a sintetizar la evidencia científica disponible sobre los modelos predictivos de caries dental basados en Big Data, inteligencia artificial y aprendizaje automático.

Este tipo de revisión permitió integrar hallazgos provenientes de estudios con diferentes diseños metodológicos, incluyendo estudios experimentales, observacionales, revisiones sistemáticas y análisis bibliométricos, con el propósito de ofrecer una visión amplia sobre los tipos de datos utilizados, los algoritmos empleados y el desempeño de estas herramientas en la predicción del riesgo de caries dental.

Se consultaron bases de datos científicas de amplio reconocimiento en ciencias de la salud y tecnología, entre ellas Scopus, PubMed, Web of Science, ScienceDirect y Google Scholar, las cuales fueron seleccionadas por su cobertura y relevancia en la difusión de literatura científica. Asimismo, se incluyeron artículos que abordaban la aplicación de inteligencia artificial en odontología, debido a su relevancia en el desarrollo de modelos predictivos.

Para la búsqueda se emplearon términos en inglés relacionados con el tema, combinados mediante operadores booleanos, tales como: “*dental caries*”, “*tooth decay*”, “*machine learning*”, “*artificial intelligence*”, “*deep learning*”, “*predictive model*” y “*big data*”. Estas palabras clave fueron utilizadas en los campos de título, resumen y palabras clave, con el objetivo de identificar estudios relevantes para la temática.

En la tabla 1 se detallaron los criterios de inclusión y exclusión considerados para la selección de los estudios analizados.

Tabla 1

Criterios de inclusión y exclusión

Criterios de inclusión	Criterios de exclusión
Estudios publicados entre 2015 y 2025.	Estudios que abordaron únicamente el diagnóstico sin componente predictivo.
Investigaciones relacionadas con modelos predictivos de caries dental.	Investigaciones no relacionadas con caries dental.
Estudios que utilizaron inteligencia artificial, aprendizaje automático o Big Data.	Cartas al editor, editoriales, opiniones o documentos sin rigor metodológico.
Artículos con metodología clara (estudios experimentales, observacionales, validación de modelos, revisiones sistemáticas y análisis bibliométricos).	Estudios duplicados o con información incompleta.
Estudios que reportaron métricas de desempeño (accuracy, AUC, sensibilidad, especificidad).	Artículos sin acceso a texto completo.
Publicaciones en idioma inglés o español.	—

La selección de la literatura se realizó mediante la revisión de títulos, resúmenes y textos completos, seguida de un análisis cualitativo de la información. Los estudios fueron organizados en categorías temáticas que incluyeron tipos de datos utilizados, algoritmos empleados, métricas de desempeño y principales hallazgos, lo que permitió identificar patrones, convergencias y limitaciones en la evidencia científica.

Los resultados fueron sistematizados en tablas que facilitaron la comparación entre estudios y la síntesis de la información, lo que permitió estructurar el análisis en torno a los enfoques metodológicos, el rendimiento de los modelos y su aplicabilidad clínica.

Dado que esta investigación se fundamentó exclusivamente en la revisión de literatura secundaria, no se requirió aprobación ética. No obstante, se garantizó en todo momento la integridad científica mediante el uso adecuado de fuentes, la correcta citación de los estudios incluidos y la presentación fiel de la información analizada.

Resultados

La tabla 2 presenta de manera estructurada los principales hallazgos de los estudios incluidos en la revisión narrativa. La información fue organizada en función de categorías temáticas previamente definidas, lo que permitió una comparación homogénea entre los tipos de estudio, los datos utilizados, los algoritmos de inteligencia artificial aplicados y las métricas de desempeño reportadas, así como la identificación de los resultados principales y las limitaciones asociadas.

Esta sistematización facilitó el análisis transversal de la evidencia disponible e hizo posible identificar patrones, convergencias y brechas en el uso de modelos predictivos de caries dental basados en inteligencia artificial, en concordancia con el objetivo del estudio.

Tabla 2

Matriz de resultados

Autor (año)	Tipo de estudio	Datos utilizados	Algoritmo utilizado	Métricas	Resultado principal
Lee et al. (2018)	Estudio evaluativo diagnóstico	3000 radiografías periapicales clasificadas por expertos	CNN (GoogLeNet Inception v3)	Precisión, sensibilidad, especificidad, VPP, VPN, ROC, AUC	Alto rendimiento diagnóstico; precisión hasta 89% y AUC 0.917
Schwendicke et al. (2019)	Revisión de alcance	Imágenes dentales (10–5166), principalmente radiografías	CNN (AlexNet, VGG, ResNet, U-Net, etc.)	Precisión, AUC, sensibilidad, especificidad, F1	Rendimiento variable; precisión en caries entre 0.82–0.89; desempeño similar o superior a humanos
Casalegno et al. (2019)	Desarrollo y validación de modelo	217 imágenes DIAGNOcam con segmentación experta	CNN tipo U-Net (autoencoder)	IoU, mIoU, ROC, AUC	mIoU 72.7%; AUC hasta 85.6% en detección de caries
Schwendicke et al. (2020)	Revisión narrativa	Datos clínicos, radiográficos y biomédicos	ML, NN, Deep Learning (CNN)	Accuracy, AUC, F1-score	IA con alto potencial clínico, limitada por calidad de datos y estandarización
Hung et al. (2019)	Estudio transversal (ML supervisado)	5135 registros NHANES con 357 variables	SVM, Random Forest, XGBoost, kNN	Precisión, sensibilidad, especificidad, AUC	SVM mejor desempeño; AUC 0.997 y precisión 97.1%
Schwendicke et al. (2020)	Estudio piloto in vitro	226 imágenes NILT de dientes extraídos	CNN (ResNet18, ResNext50)	AUC, sensibilidad, especificidad, PPV, NPV	Rendimiento aceptable (AUC 0.74); útil como apoyo diagnóstico
García et al. (2020)	Estudio de precisión diagnóstica	3686 radiografías bitewing	CNN tipo U-Net	Accuracy, sensibilidad, especificidad, F1, MCC, IoU	CNN superior a dentistas; mayor sensibilidad (0.75 vs 0.36)
Dettori et al. (2025)	Revisión bibliométrica	50 artículos (PubMed, Scopus, Embase)	VOSviewer (análisis de redes)	Análisis de clusters	IA mejora diagnóstico y personalización; énfasis en ética y explicabilidad
Park & Park (2018)	Revisión	Datos clínicos, conductuales e imágenes	ANN, ML, redes bayesianas	Precisión	IA mejora diagnóstico y decisiones clínicas; precisión hasta 100% en algunos modelos
Krois et al. (2019)	Estudio experimental	2001 segmentos de radiografías panorámicas	CNN profunda	Accuracy, AUC, F1, sensibilidad, especificidad	Rendimiento similar a dentistas; precisión ~0.81
Khanagar et al. (2020)	Revisión sistemática	43 estudios de múltiples bases de datos	CNN, ANN, redes bayesianas	Accuracy, AUC, ROC, sensibilidad, especificidad	IA con rendimiento excelente; en algunos casos superior a especialistas

En relación con los algoritmos empleados, como se muestra en la tabla 3, los más utilizados en los estudios analizados correspondieron principalmente a las redes neuronales convolucionales (CNN)

dentro del enfoque de aprendizaje profundo, aplicadas predominantemente al análisis de imágenes radiográficas dentales.

Estos modelos evidenciaron un alto desempeño diagnóstico, con valores de precisión entre 0.80 y 0.89 y AUC entre 0.74 y 0.92, alcanzando en varios casos un rendimiento comparable o superior al de los profesionales odontológicos.

Asimismo, las variantes de CNN orientadas a la segmentación de imágenes, como las arquitecturas tipo U-Net, mostraron resultados igualmente elevados, destacándose en la detección y localización de lesiones, con valores de mIoU cercanos al 72.7% y AUC entre 0.83 y 0.85.

Por otra parte, los modelos de aprendizaje automático supervisado, como la máquina de soporte vectorial (SVM), presentaron el mayor desempeño reportado, con una precisión del 97.1% y un AUC de 0.997, seguidos por algoritmos como Random Forest, XGBoost, kNN y regresión logística, que también mostraron altos niveles de rendimiento predictivo.

En cuanto a las redes neuronales artificiales (ANN) y los modelos híbridos que combinan diferentes enfoques de inteligencia artificial, se observó un desempeño igualmente elevado, aunque con variabilidad en función del tipo de datos y del diseño metodológico de los estudios.

Finalmente, los estudios de tipo revisión evidenciaron un alto potencial de la inteligencia artificial en odontología, aunque carecieron de validación experimental directa, lo que limita la interpretación de sus resultados. En conjunto, los hallazgos reflejan una tendencia creciente hacia el uso de modelos de aprendizaje profundo, especialmente en tareas basadas en imágenes, con resultados prometedores en el diagnóstico automatizado y la predicción del riesgo de caries dental.

Tabla 3

Síntesis de algoritmos y desempeño en los estudios incluidos

Algoritmo	Tipo de aprendizaje	Estudios que lo utilizan	Tipo de datos	Métrica promedio	Desempeño general
CNN (Redes Neuronales Convolucionales: U-Net, ResNet, GoogLeNet, VGG)	Aprendizaje profundo (Deep Learning)	Lee et al., 2018; Schwendicke et al., 2019; Casalegno et al., 2019; Schwendicke et al., 2020 (in vitro); Garcia et al., 2020; Krois et al., 2019; Khanagar et al., 2020	Imágenes radiográficas (periapicales, bitewing, panorámicas), NILT, DIAGNOcam	Accuracy: 0.80–0.89 / AUC: 0.74–0.92	Alto (rendimiento similar o superior a dentistas en varios estudios)
CNN + variantes (autoencoder, segmentación tipo U-Net)	Aprendizaje profundo	Casalegno et al., 2019; Garcia et al., 2020	Imágenes con segmentación (pixel a pixel)	mIoU: ~72.7% / AUC: 0.83–0.85	Alto (buena detección y localización de lesiones)
SVM (Support Vector Machine)	Aprendizaje automático supervisado	Hung et al., 2019	Datos clínicos y epidemiológicos (NHANES)	Accuracy: 97.1% / AUC: 0.997	Muy alto (mejor modelo del estudio)
Random Forest, XGBoost, kNN, Regresión Logística	Aprendizaje automático supervisado	Hung et al., 2019	Datos clínicos estructurados	Accuracy: >90%	Alto (ligeramente inferior a SVM)

Redes neuronales artificiales (ANN)	Aprendizaje automático / profundo	Park & Park, 2018; Khanagar et al., 2020	Datos clínicos, imágenes, registros digitales	Accuracy: 80%–100%	Alto (especialmente en modelos entrenados)
Modelos múltiples (CNN, ANN, BN, PNN)	Mixto (ML + DL)	Khanagar et al., 2020	Datos clínicos e imágenes	Variable según estudio	Alto (rendimiento comparable a expertos)
IA sin implementación directa (revisiones)	No aplica	Schwendicke et al., 2020 (revisión narrativa); Dettori et al., 2025	Datos conceptuales, clínicos y bibliométricos	No reporta métricas directas	Evidencia potencial alta, sin validación experimental directa

En relación con los tipos de datos, la tabla 4 muestra que predominaron las imágenes radiográficas dentales como principal fuente de información, siendo utilizadas en más de la mitad de los estudios analizados. Esto se explica por su alta utilidad en el diagnóstico automatizado, especialmente en la detección de caries y pérdida ósea periodontal.

En menor proporción, se emplearon datos clínicos estructurados y datos mixtos, los cuales contribuyen a fortalecer la capacidad predictiva de los modelos al integrar múltiples variables clínicas y contextuales.

Por otra parte, los datos de imagen avanzados, como los obtenidos mediante tecnologías de transiluminación o sistemas especializados, mostraron una mayor precisión en la localización de lesiones; sin embargo, su uso se vio limitado por la disponibilidad de los equipos y el tamaño reducido de las muestras.

Finalmente, los estudios basados en datos bibliométricos aportaron una visión general sobre las tendencias de investigación en inteligencia artificial en odontología, aunque no proporcionaron evidencia clínica directa sobre el desempeño diagnóstico.

Tabla 4

Tipos de datos utilizados en los estudios incluidos

Tipo de datos	Estudios que lo utilizan	Frecuencia de uso (%)	Ventajas	Limitaciones
Imágenes radiográficas dentales (periapicales, bitewing, panorámicas, CBCT)	Lee et al., 2018; Schwendicke et al., 2019; Schwendicke et al., 2020 (in vitro); Garcia et al., 2020; Krois et al., 2019; Khanagar et al., 2020	~55%	Alta precisión diagnóstica; permiten automatizar detección de caries y pérdida ósea; ampliamente disponibles	Requieren grandes volúmenes de datos; variabilidad en calidad de imagen; dependencia de anotación experta
Imágenes avanzadas y segmentadas (DIAGNOcam, NILT)	Casalegno et al., 2019; Schwendicke et al., 2020 (in vitro)	~18%	Permiten análisis más detallado (pixel a pixel); mejor localización de lesiones	Tamaño de muestra reducido; menor disponibilidad clínica; mayor complejidad técnica
Datos clínicos y epidemiológicos estructurados	Hung et al., 2019; Park & Park, 2018	~18%	Alta precisión predictiva; fáciles de procesar con ML; útiles para predicción de riesgo	No capturan información visual; dependen de calidad del registro clínico

Datos mixtos (imágenes + clínicos + registros digitales)	Schwendicke et al., 2020 (revisión); Khanagar et al., 2020	~18%	Enfoque integral; mayor capacidad predictiva y personalización	Integración compleja; problemas de estandarización y acceso a datos
Datos bibliométricos y metadatos	Dettori et al., 2025	~9%	Permiten analizar tendencias y evolución científica; identifican áreas de investigación	No aportan evidencia clínica directa; no evalúan desempeño diagnóstico

En cuanto a las métricas de evaluación, como se observa en la tabla 5, las más empleadas fueron la precisión (accuracy), sensibilidad, especificidad y el área bajo la curva (AUC), lo que evidencia un enfoque predominante en la evaluación del rendimiento diagnóstico de los modelos. El AUC se consolidó como uno de los indicadores más robustos, con valores que oscilaron entre 0.74 y 0.997, lo que refleja una elevada capacidad discriminativa en diversos estudios.

Asimismo, métricas como el Intersection over Union (IoU) y el coeficiente Dice fueron utilizadas principalmente en tareas de segmentación de imágenes, mientras que el F1-score y el coeficiente de correlación de Matthews (MCC) resultaron especialmente útiles en contextos con datos desbalanceados. En conjunto, se evidenció una alta heterogeneidad en las métricas reportadas, lo que dificulta la comparación directa entre los estudios analizados.

Tabla 5

Métricas de desempeño y resultados

Métrica	Estudios que la reportan	Rango de valores	Interpretación
Accuracy (Precisión)	Lee et al., 2018; Schwendicke et al., 2019; Hung et al., 2019; Garcia et al., 2020; Krois et al., 2019; Park & Park, 2018; Khanagar et al., 2020	0.77 – 0.97	Indica el rendimiento global del modelo; valores altos reflejan buena capacidad diagnóstica
AUC (Área bajo la curva ROC)	Lee et al., 2018; Schwendicke et al., 2019; Casalegno et al., 2019; Hung et al., 2019; Schwendicke et al., 2020 (in vitro); Khanagar et al., 2020	0.74 – 0.997	Mide la capacidad discriminativa; valores cercanos a 1 indican excelente desempeño
Sensibilidad (Sensitivity)	Lee et al., 2018; Schwendicke et al., 2019; Hung et al., 2019; Schwendicke et al., 2020 (in vitro); Garcia et al., 2020; Krois et al., 2019; Khanagar et al., 2020	0.59 – 0.996	Evalúa la capacidad de detectar correctamente casos positivos (caries)
Especificidad (Specificity)	Lee et al., 2018; Schwendicke et al., 2019; Hung et al., 2019; Schwendicke et al., 2020 (in vitro); Garcia et al., 2020; Krois et al., 2019; Khanagar et al., 2020	0.63 – 0.94	Mide la capacidad de identificar correctamente casos negativos
VPP / PPV	Lee et al., 2018; Hung et al., 2019; Schwendicke et al., 2020 (in vitro); Garcia et al., 2020; Krois et al., 2019; Khanagar et al., 2020	0.63 – 0.95	Probabilidad de que un resultado positivo sea correcto
VPN / NPV	Lee et al., 2018; Schwendicke et al., 2020 (in vitro); Garcia et al., 2020; Krois et al., 2019; Khanagar et al., 2020	0.73 – 0.80	Probabilidad de que un resultado negativo sea correcto
IoU / mIoU	Casalegno et al., 2019; Garcia et al., 2020	0.49 – 0.72	Evalúa la precisión en segmentación de imágenes (nivel píxel)
F1-score	Schwendicke et al., 2019; Garcia et al., 2020; Krois et al., 2019	No homogéneo	Balance entre precisión y sensibilidad; útil en datos desbalanceados

MCC (Coeficiente de correlación de Matthews)	García et al., 2020	No reportado en rango común	Métrica robusta para clasificación binaria, incluso con clases desbalanceadas
Dice coefficient	Schwendicke et al., 2019	No homogéneo	Evalúa similitud en segmentación de imágenes
ICC (Coeficiente de correlación intraclase)	Khanagar et al., 2020	No homogéneo	Mide concordancia entre evaluadores o modelos

Finalmente, como se sintetiza en la tabla 6, las principales limitaciones identificadas incluyeron el tamaño reducido de las muestras, la heterogeneidad en los datos y la falta de estandarización en las métricas, las cuales se presentaron con alta frecuencia y afectaron la comparabilidad entre estudios. Asimismo, la limitada validación externa y el uso predominante de datos retrospectivos restringieron la generalización de los hallazgos.

Otras limitaciones relevantes fueron el desbalance de datos, la variabilidad en la anotación de expertos y la falta de replicabilidad de los modelos, factores que afectan directamente su aplicabilidad clínica. En conjunto, estas limitaciones evidencian la necesidad de desarrollar estudios más robustos, con metodologías estandarizadas y validación en entornos clínicos reales, en línea con los objetivos planteados en esta revisión.

Tabla 6

Limitaciones comunes

Tipo de limitación	Frecuencia	Impacto en resultados
Tamaño de muestra reducido	Alta	Limita la generalización de los modelos y puede generar sobreajuste, reduciendo la validez externa.
Heterogeneidad en los datos (tipo de imágenes, fuentes, variables clínicas)	Alta	Dificulta la comparación entre estudios y afecta la consistencia de los resultados.
Falta de estandarización en métricas	Alta	Impide la comparación directa del rendimiento entre modelos y estudios.
Uso de datos retrospectivos o secundarios	Moderada	Puede introducir sesgos y limitar la calidad de la evidencia.
Limitada validación externa	Alta	Reduce la aplicabilidad clínica en diferentes poblaciones o contextos.
Desbalance en los datos	Moderada	Afecta la precisión de métricas como accuracy, favoreciendo clases mayoritarias.
Dependencia de imágenes 2D (radiografías)	Alta	Puede limitar la detección precisa frente a métodos más avanzados como CBCT o NILT.
Variabilidad en la anotación por expertos	Moderada	Introduce sesgos en el entrenamiento y evaluación de los modelos.
Falta de transparencia y replicabilidad de los modelos (IA)	Moderada	Dificulta la interpretación clínica y la adopción en la práctica odontológica.

Limitaciones tecnológicas y de acceso a datos	Moderada	Restringe el desarrollo y entrenamiento de modelos más robustos.
Diseños de estudio in vitro o piloto	Baja-Moderada	Reduce la aplicabilidad en escenarios clínicos reales.
Ausencia de comparación con estándar clínico sólido	Moderada	Limita la validez de los resultados frente a la práctica real.

Discusión

Los hallazgos de la presente revisión evidenciaron que la aplicación de técnicas de inteligencia artificial y aprendizaje automático en la predicción de caries dental ha alcanzado niveles elevados de desempeño, particularmente en modelos basados en imágenes y algoritmos de aprendizaje profundo.

En este contexto, las métricas de evaluación reportadas (tabla 5) permitieron analizar de manera crítica tanto la capacidad diagnóstica y predictiva de estos modelos como sus principales limitaciones metodológicas, aportando evidencia relevante en relación con el objetivo del estudio.

En primer lugar, el área bajo la curva (AUC) se posicionó como uno de los indicadores más robustos para evaluar el rendimiento de los modelos, con valores que oscilaron entre 0.74 y 0.997 (tabla 5). Estos resultados sugieren una elevada capacidad discriminativa en la mayoría de los estudios, especialmente en aquellos que emplearon algoritmos como SVM y redes neuronales convolucionales, como reportaron Hung et al. (2019) y Lee et al. (2018).

Este hallazgo es consistente con lo señalado por Schwendicke et al. (2019), quienes destacan el AUC como un indicador clave en la evaluación de modelos diagnósticos en odontología, lo que refuerza la validez de los resultados obtenidos.

Asimismo, la precisión (accuracy), que varió entre 0.77 y 0.97 (tabla 5), indicó un alto rendimiento global en la clasificación de casos con y sin caries. Estudios como los de Garcia et al. (2020) y Krois et al. (2019) reportaron niveles de precisión comparables a los de profesionales odontológicos, lo que sugiere el potencial de estas herramientas como sistemas de apoyo clínico.

Sin embargo, esta métrica debe interpretarse con cautela, ya que puede verse afectada por el desbalance en los datos, tal como advierten Hung et al. (2019). En este sentido, métricas complementarias como el F1-score y el coeficiente de correlación de Matthews (MCC), reportadas por Garcia et al. (2020), adquieren mayor relevancia al proporcionar una evaluación más equilibrada del rendimiento en contextos con clases desiguales.

Por otra parte, la sensibilidad y la especificidad, con rangos de 0.59 a 0.996 y 0.63 a 0.94 respectivamente (tabla 5), reflejaron la capacidad de los modelos para identificar correctamente tanto casos positivos como negativos. La alta sensibilidad observada en estudios como Lee et al. (2018) y Khanagar et al. (2020) resulta especialmente relevante desde el punto de vista clínico, ya que favorece la detección temprana de lesiones cariosas, un aspecto clave en la prevención.

No obstante, la variabilidad en estos valores, también señalada por Schwendicke et al. (2019), indica que el desempeño de los modelos depende en gran medida del tipo de datos utilizados, la calidad de estos y el algoritmo implementado, lo que limita su generalización. En relación con las tareas de segmentación de imágenes, métricas como el Intersection over Union (IoU) y el coeficiente Dice demostraron ser fundamentales para evaluar la precisión a nivel de píxel (tabla 5).

Casalegno et al. (2019) y Garcia et al. (2020) reportaron valores de IoU entre 0.49 y 0.72, lo que sugiere un desempeño adecuado, aunque todavía limitado en la delimitación precisa de las lesiones. Este resultado coincide con lo descrito por Schwendicke et al. (2020), quienes señalan que la segmentación exacta continúa siendo un desafío técnico relevante para la implementación clínica de la inteligencia artificial.

A pesar de estos resultados prometedores, la alta heterogeneidad en las métricas reportadas (tabla 5) constituye una limitación crítica en la literatura actual. Diferentes estudios, como los de Khanagar et al. (2020) y Dettori et al. (2025), destacan la ausencia de estandarización en los métodos de evaluación, lo que dificulta la comparación directa entre modelos y limita la reproducibilidad y generalización de los hallazgos.

Esta variabilidad metodológica evidencia la necesidad de establecer criterios unificados para la evaluación del desempeño de modelos predictivos en odontología. Además, aunque los modelos presentan un alto rendimiento en términos métricos, su aplicabilidad clínica permanece condicionada por múltiples factores estructurales y metodológicos.

Entre ellos, destacan la falta de validación externa y el uso predominante de datos retrospectivos, tal como señalan Schwendicke et al. (2020) y Park & Park (2018), lo que reduce la validez externa y la transferibilidad de los modelos a contextos clínicos reales. Asimismo, la limitada interpretabilidad de los modelos, destacada por Dettori et al. (2025), representa una barrera significativa para su adopción en la práctica clínica, donde la transparencia en la toma de decisiones es un requisito fundamental.

En conjunto, los resultados analizados indican que las herramientas basadas en inteligencia artificial poseen un alto potencial para mejorar la predicción y el diagnóstico de caries dental, en concordancia con lo planteado por Wang et al. (2025) y Ding et al. (2023). No obstante, para lograr su implementación efectiva en la práctica clínica, es imprescindible fortalecer el rigor metodológico, estandarizar las métricas de evaluación y promover el desarrollo de estudios con diseños prospectivos y validación externa.

Finalmente, estos avances deben ir acompañados de estrategias orientadas a mejorar la calidad y disponibilidad de los datos, así como a incrementar la explicabilidad de los modelos, con el fin de facilitar su integración en la toma de decisiones clínicas. De este modo, será posible avanzar hacia una odontología más predictiva, personalizada y basada en evidencia, alineada con las tendencias actuales de la salud digital.

Conclusiones

La evidencia sintetizada en esta revisión permite concluir que la aplicación de técnicas de Big Data, inteligencia artificial y aprendizaje automático en la predicción del riesgo de caries dental constituye una herramienta eficaz, con alto potencial clínico para mejorar la detección temprana y optimizar la toma de decisiones en odontología.

De manera consistente, los modelos basados en redes neuronales convolucionales (CNN), máquinas de vectores de soporte (SVM) y redes neuronales artificiales (ANN) demostraron un elevado rendimiento diagnóstico, con valores altos de precisión, sensibilidad y AUC (tabla 5), lo que respalda su utilidad como herramientas complementarias e incluso superiores en determinados contextos frente a los métodos tradicionales de evaluación del riesgo de caries.

Asimismo, el uso predominante de imágenes radiográficas y la integración de datos clínicos han contribuido de manera significativa a fortalecer la capacidad predictiva de estos modelos, consolidando

el papel de la inteligencia artificial como una herramienta innovadora y en expansión dentro de la odontología contemporánea.

En relación con los factores que condicionan el desempeño de estos modelos, se identificó que su efectividad está determinada por múltiples variables metodológicas, técnicas y contextuales, las cuales impactan directamente en la validez y generalización de los resultados.

Aspectos como el tamaño de la muestra, la calidad y heterogeneidad de los datos, el desbalance en las clases, la variabilidad en la anotación de expertos y la falta de estandarización en las métricas de evaluación (tabla 5) emergieron como limitaciones críticas.

Asimismo, la dependencia de datos retrospectivos y la limitada validación externa reducen la aplicabilidad clínica de los modelos en diferentes poblaciones. Estos hallazgos sugieren que el rendimiento de los modelos no depende exclusivamente del algoritmo empleado, sino de la calidad del diseño metodológico, la representatividad de los datos y la adecuación del modelo al contexto clínico, lo que resalta la necesidad de desarrollar enfoques más robustos, estandarizados y adaptados a escenarios reales.

Respecto a la implementación clínica, se identificaron oportunidades relevantes para integrar estas tecnologías en la práctica odontológica, especialmente en el apoyo al diagnóstico, la predicción del riesgo y la planificación de tratamientos personalizados.

Sin embargo, persisten barreras significativas, entre ellas la falta de replicabilidad de los modelos, las limitaciones tecnológicas, el acceso restringido a grandes volúmenes de datos de calidad y la ausencia de protocolos estandarizados para su uso clínico.

En este sentido, resulta fundamental promover el desarrollo de modelos más interpretables, realizar validaciones externas en entornos clínicos reales e incentivar la integración de datos multimodales, con el fin de mejorar la precisión, confiabilidad y aplicabilidad de estos sistemas.

En conjunto, estos avances podrían facilitar la adopción progresiva de la inteligencia artificial en odontología, contribuyendo a la transición hacia una práctica más preventiva, personalizada y basada en evidencia, en consonancia con los principios de la salud digital y la medicina de precisión.

Referencias

- Casalegno, F., Newton, T., Daher, R., Abdelaziz, M., Lodi-Rizzini, A., Schürmann, F., Krejci, I., & Markram, H. (2019). Caries detection with near-infrared transillumination using deep learning. *Journal of Dental Research*, 98(11), 1227–1233. <https://doi.org/10.1177/0022034519871884>
- Dettori, M., Lamoum, D., Lingström, P., & Campus, G. (2025). Artificial intelligence and innovation in oral health care sciences: A conceptual review. *Healthcare*, 13(24), 3327. <https://doi.org/10.3390/healthcare13243327>
- Ding, H., Wu, J., Zhao, W., Matinlinna, J., Burrow, M., & Tsoi, J. (2023). Artificial intelligence in dentistry—A review. *Frontiers in Dental Medicine*, 4, 1085251. <https://doi.org/10.3389/fdmed.2023.1085251>
- Garcia, A., Gehrung, S., Krois, J., Chaurasia, A., Rossi, J., Gaudin, R., Elhennawy, K., & Schwendicke, F. (2020). Detecting caries lesions of different radiographic extension on bitewings using deep learning. *Journal of Dentistry*, 100, 103425. <https://doi.org/10.1016/j.jdent.2020.103425>

- Havsted, K., Hänsel, G., Isberg, P., Pigg, M., Svensäter, G., Foresight Research Consortium, & Rohlin, M. (2023). Multivariable prediction models of caries increment: A systematic review and critical appraisal. *Systematic Reviews*, 12, 202. <https://doi.org/10.1186/s13643-023-02298-y>
- Hung, M., Voss, M., Rosales, M., Li, W., Su, W., Xu, J., Bounsanga, J., Ruiz-Negrón, B., Lauren, E., & Licari, F. (2019). Application of machine learning for diagnostic prediction of root caries. *Gerodontology*, 36(4), 395–404. <https://doi.org/10.1111/ger.12432>
- Khanagar, S., Al-Ehaideb, A., Maganur, P., Vishwanathaiah, S., Patil, S., Baeshen, H., Sarode, S., & Bhandi, S. (2021). Developments, application, and performance of artificial intelligence in dentistry – A systematic review. *Journal of Dental Sciences*, 16(1), 508–522. <https://doi.org/10.1016/j.jds.2020.06.019>
- Krois, J., Ekert, T., Meinhold, L., Golla, T., Kharbot, B., Wittemeier, A., Dörfer, C., & Schwendicke (2019). Deep learning for the radiographic detection of periodontal bone loss. *Scientific Reports*, 9, 8495. <https://doi.org/10.1038/s41598-019-44839-3>
- Lee, J., Kim, D., Jeong, S., & Choi, S. (2018). Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *Journal of Dentistry*, 77, 106–111. <https://doi.org/10.1016/j.jdent.2018.07.015>
- Park, W., & Park, J. (2018). History and application of artificial neural networks in dentistry. *European Journal of Dentistry*, 12(4), 594–601. https://doi.org/10.4103/ejd.ejd_325_18
- Schwendicke, F., Elhennawy, K., Paris, S., Friebertshäuser, P., & Krois, J. (2020). Deep learning for caries lesion detection in near-infrared light transillumination images: A pilot study. *Journal of Dentistry*, 92, 103260. <https://doi.org/10.1016/j.jdent.2019.103260>
- Schwendicke, F., Golla, T., Dreher, M., & Krois, J. (2019). Convolutional neural networks for dental image diagnostics: A scoping review. *Journal of Dentistry*, 91, 103226. <https://doi.org/10.1016/j.jdent.2019.103226>
- Schwendicke, F., Samek, W., & Krois, J. (2020). Artificial intelligence in dentistry: Chances and challenges. *Journal of Dental Research*, 99(7), 769–774. <https://doi.org/10.1177/0022034520915714>
- Shah, S. (2023). Caries and public health. *Journal of Khyber College of Dentistry*, 12(3). <https://doi.org/10.33279/jkcd.v12i03.74>
- Wang, L., Xu, Y., Wang, W., & Lu, Y. (2025). Application of machine learning in dentistry: Insights, prospects and challenges. *Applied Oral Sciences*, 84, 145–154. <https://journalofkcd.com/kcd/article/view/74>
- Wang, X., Zhang, P., Lu, H., Luo, D., Yang, D., Li, K., Qiu, S., Zeng, H., & Zeng, X. (2025). Risk prediction models for dental caries in children and adolescents: A systematic review and meta-analysis. *BMJ Open*, 15(3), e088253. <https://doi.org/10.1136/bmjopen-2024-088253>
- World Health Organization. (2025). *Sugars and dental caries*. <https://www.who.int/news-room/fact-sheets/detail/sugars-and-dental-caries>

Declaraciones éticas

Conflicto de interés

La autora declara que no existe ningún tipo de conflicto de interés relacionado con la presente investigación.

Fuente de financiamiento

La investigación fue financiada en su totalidad por la propia autora.

Contribución de autoría

Noelia Amaray Velastegui Almeida: Conceptualización, metodología, análisis formal, investigación, gestión de datos, redacción - preparación del borrador original, redacción - revisión y edición, financiamiento, administración del proyecto, recursos, supervisión.

Gonzalo Paúl Rodríguez Galarza: Conceptualización, metodología, software, validación, análisis formal, investigación, gestión de datos, visualización, redacción - preparación del borrador original, redacción - revisión y edición, financiamiento, recursos.

Los autores intervinieron de manera activa en el análisis de los resultados, así como en la revisión crítica y la aprobación de la versión final del manuscrito para su publicación.